

Eine Methode zur Strukturanpassung von Umfragedaten

Kopp, Bruno

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Kopp, B. (2006). Eine Methode zur Strukturanpassung von Umfragedaten. *ZUMA Nachrichten*, 30(59), 89-110. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-69267>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:
<https://creativecommons.org/licenses/by-nc-nd/4.0>

EINE METHODE ZUR STRUKTURANPASSUNG VON UMFRAGEDATEN

A METHOD FOR STRUCTURAL FITTING IN SURVEY SAMPLING

BRUNO KOPP

Es wird ein Verfahren zur Anpassung von Befragungen an Sollvorgaben vorgestellt. Dabei wird jedem Befragten i ein Gewicht der Form $f(\gamma_i)$ derart zugeordnet, dass die Strukturen hinsichtlich vorgegebener Befragungsmerkmale mit den Sollwerten dieser Merkmale überein stimmen, wobei die Faktoren $f(\gamma_i)$ möglichst nahe an vorgegebenen Eingangsfaktoren t_i liegen. Durch geeignete Wahl der Funktion f lassen sich Bedingungen nach Vorzeichen und Lage an die Personengewichte stellen. Ein Vergleich mit anderen Algorithmen zeigt die besondere Robustheit der Methode.

The paper presents a calibration method which assigns individuals i of a survey a weight $f(\gamma_i)$ in order to fit the structure of known population totals. The procedure generates weights which are close to a set of given initial factors t_i . An appropriate selection of the function f allows different properties of the generated weights in sign and range. A comparison with other methods reveals the stability of the algorithm.

1 Problemstellung

Gewichtungsverfahren werden benutzt zur Anpassung von Umfragedaten an vorgegebene Sollstrukturen (etwa aus der amtlichen Statistik).

Wir können – ohne Einschränkung der Allgemeinheit – die Strukturmerkmale hier allgemein als binäre Vektoren annehmen: Eine ursprüngliche Variable, etwa das Alter mit den Kategorien 1 bis 3, wird dabei aufgelöst in 3 dichotome Merkmale. Die Ursprungsvariable kann ein zusammengesetztes Merkmal (etwa Alter * Bildung) sein, so dass beliebig komplexe Strukturen erfasst werden. Ein solches Merkmal unterscheidet sich prinzipiell nicht von einem einfachen Merkmal; insofern ist auch die Redewendung „Anpassung an Randverteilungen“ eher künstlich.

Ist r die Gesamtanzahl der Kategorien aller ausgewählten Anpassungsmerkmale und n die Befragtenanzahl der Umfrage, so lassen sich die Anpassungsmerkmale in einer binären

Matrix $S = (s_{ik})$ mit r Zeilen und n Spalten eindeutig festlegen. Gesucht werden nun individuelle Faktoren g_i für jeden der n Befragten derart, dass für jede Kategorie gilt

$$\sum_{k=1}^n s_{ik} g_k = b_i, \quad (i=1, \dots, r).$$

Dabei sind die b_i vorgegebene Sollwerte, also Gesamtfallzahlen der Kategorie i als Zielgrößen für die Umfrage.

Sind die Daten der Umfrage – also die Matrix S – und die Sollwerte in sich widerspruchsfrei, so kann man die Gewichtungsfaktoren in der Lösungsmenge des inhomogenen linearen Gleichungssystems (System der Restriktionen)

$$Sg = b \quad (1)$$

suchen.

Dabei ist $g = (g_1, \dots, g_n)^T$ und $b = (b_1, \dots, b_r)^T$ mit T als Transpositionszeichen. Da die Lösung von (1) – wenn sie existiert – nicht eindeutig (weil im Normalfall $r < n$) ist, wird in der Regel die Zusatzforderung gestellt, dass die Gewichte g_i von Vorgewichten t_i möglichst wenig abweichen sollten. Unter Vorgewichten verstehen wir hier so genannte Designgewichte.

„Mit der Designgewichtung wird verhindert, dass eine vom Ansatz her bewusste (z.B. ein überhöhter Stichprobenanteil in den östlichen Bundesländern) oder nicht zu verhindernde (z.B. bei haushaltsbasierten Bevölkerungsstichproben) Verzerrung der Abbildung in die Auswertung übertragen wird.“ (Von der Heyde 1999).

Eine sinnvolle Forderung ist, dass die Gewichte positiv sein müssen, was bei einigen Methoden allerdings mit zunehmender Anzahl r an Restriktionen zu einem besonderen Problem wird.

2 Der Iterative Proportional Fitting – Algorithmus („IPF“)

Diese auf Deming & Stephan (1940) zurückgehende Gewichtungsmethode ermittelt die Gewichte als Quotienten aus Fallzahlen der Sollzellen zu den entsprechenden Fallzahlen der Umfrage (Ist-Zellen). Auf Basis dieser Idee arbeiten in der Praxis viele Methoden, deren genaue Beschreibungen leider öffentlich nicht zugänglich sind. Wir müssen uns daher auf eine prinzipielle Erörterung des IPF beschränken, dessen Arbeitsweise an folgendem Beispiel verdeutlicht werden soll. Alle Männer erhalten zunächst als Gewicht den

Quotienten Anzahl Männer (Soll) zu der Anzahl Männer (Ist), wobei bei der Berechnung der Ist-Werte die Vorgewichte benutzt werden. Entsprechend erhalten die Frauen ihr Gewicht. Der erhaltene Gewichtsvektor enthält die vorläufigen Gewichte (1. Stufe). Anschließend wird mit einem weiteren Merkmal (Z.B. Alter) ganz entsprechend verfahren, wobei die Fälle in den Ist-Zellen nun mit den bereits vorliegenden Gewichten der Stufe 1 ermittelt werden. Das ergibt wieder Gewichte, die – mit den Gewichten der 1. Stufe multipliziert – die neuen Gewichte der Stufe 2 liefern. Das Verfahren wird über alle vorgesehenen Anpassungsmerkmale fortgesetzt. Dabei werden stets die gerade erzeugten Gewichte mit denjenigen der Vorstufe multipliziert.

Nachdem so alle Merkmale behandelt wurden, wird der Prozess wiederholt bis nach einigen Iterationen Konvergenz eintritt. Die Vorteile des IPF liegen in seiner einfachen Arbeitsweise und in dem Umstand, dass die gefundenen Faktoren aufgrund der Quotientenbildung nicht negativ sein können. Nachteile ergeben sich dann, wenn leere Ist-Zellen auftreten (Division durch 0) – was eine umständliche Zusammenlegung von Kategorien erforderlich macht. Außerdem können bei der Forderung von Unter- und Obergrenzen der Gewichte (IPF neigt zu extremen Gewichten) Konvergenzprobleme auftreten.

Näheres über IPF findet man bei Rösch (1994).

3 Abstandsminimierung zu Vorgewichten bei linearen Restriktionen („MinVar“)

Es gibt eine Reihe von Methoden, die aus der Lösungsmenge des Gleichungssystems der Restriktionen diejenige ermitteln, deren Abstand zu den Vorgewichten minimal ist.

Diese Verfahren unterscheiden sich hauptsächlich in dem verwendeten Abstandsbegriff.

So benutzt Merz (1993, 2004) ein informationstheoretisches Distanzmaß, wodurch negative Gewichte verhindert werden.

Mit dem Quadratischen Euklidischen Abstand zwischen den Gewichten und den Vorgewichten liefert die Minimierungsaufgabe mit $t = (t_1, \dots, t_n)^T$

$$\text{Min } (g-t)^T(g-t) \text{ unter der Nebenbedingung } S g = b \quad (2)$$

die Lösung

$$g = t + S^T (SS^T)^{-1} [b - S t]. \quad (3)$$

Obige Lösung ergibt sich per Lagrange-Technik (siehe Anhang).

Rösch (1994) bezeichnet dieses Verfahren als MinVar - Methode, weil im Falle gleicher Eingangsgewichte sich eine Lösung minimaler Varianz ergibt. Nachteilig bei dieser Methode ist einerseits der Umstand, dass obige Produkt-Matrix nicht invertierbar sein kann und andererseits negative Gewichte nicht auszuschließen sind.

Hinweise zum Rang der Matrizen S und SS^T

Die Inversion in (3) setzt voraus, dass S den maximalen Rang r hat, also keine der r Zeilen von S sich aus den übrigen linear bilden lässt – es dürfen also keine redundanten Informationen in S enthalten sein.

Fehlt eine Ausprägung (Missing Value, leere Zelle), führt dies zu einer Nullzeile in S , was den Rang von S reduziert.

Selbst dann, wenn alle Kategorien vorkommen, ergeben sich Redundanzen, wie folgendes Beispiel zeigt.

Entsprechen z.B. – wie eingangs erwähnt – die ersten 3 Zeilen den Kategorien des Alters, so muss die Summe der 3 ersten Zeilen in S ausschließlich zu einer Zeile aus Einsen führen.

Werden bei dem nächsten Ursprungsmerkmal – Beispiel Bildung mit 4 Kategorien – sämtliche 4 binäre Einzelmerkmale in S angefügt, so kann S automatisch nicht mehr den maximalen Rang haben. Denn die Summe dieser 4 Merkmale ergibt ja wieder lauter Einsen. Um diesen Effekt zu vermeiden, sollte man nur bei einem einzigen Ursprungsmerkmal (etwa beim Alter) sämtliche Kategorien berücksichtigen und bei allen nachfolgenden jeweils eine Ausprägung zur Vermeidung von Redundanzen auslassen.

Besondere Probleme ergeben sich bei Kombinationsmerkmalen (etwa bei der Variablen Alter * Bildung), insbesondere dann, wenn Redundanzen aus Merkmalen und deren Kombinationen (etwa die 3 Merkmale Alter, Bildung und Alter * Bildung) vorgesehen sind.

Hier könnte der Einsatz der Pseudoinversen Abhilfe schaffen, worunter dann allerdings die Genauigkeit der Ergebnisse leiden wird. Siehe Anhang 7.2.

Selbst dann, wenn keine Redundanzen vorkommen, kann der Rang von S kleiner als r sein.

Dies ist etwa der Fall, wenn zwei Kategorien verschiedener Ursprungsmerkmale ein identisches Muster aufweisen. Bei fast identischen Mustern ist die Korrelation zwischen beiden Zeilen von S nahe 1; die Matrix SS^T ist „schlecht konditioniert“, und die Invertierung kann im günstigsten Fall mittels Pseudoinverse erfolgen, wobei es oft sinnvoller wäre, eine Anpassungsvariable zu entfernen. Siehe auch die Hinweise im Anhang 7.6.

All diese Probleme spielen bei dem nachfolgenden Modell keine Rolle mehr.

4 Alternativmodell: Minimierung ohne Nebenbedingen („Min I“)

Als Alternative zur obigen Minimierungsaufgabe (2) mit linearen Restriktionen bietet sich der folgende Ansatz einer Minimierung ohne Nebenbedingungen an.

$$Z = \lambda(Sg - b)^T(Sg - b) + (g - t)^T(g - t) = \text{Min} ! \quad (4)$$

Hier ist λ ein positiver Kontrollparameter, der eine differenzierte Gewichtung beider Summanden ermöglicht. Hohe Werte für λ sorgen dafür, dass die Restriktionen (1) genauer eingehalten werden. Geht λ gegen 0, so nähern sich die Gewichte den Vorgewichten.

Die Lösung dieser quadratischen Minimierungsaufgabe (Herleitung im Anhang) ist

$$g = t + S^T(E_r / \lambda + SS^T)^{-1} [b - St]. \quad (5)$$

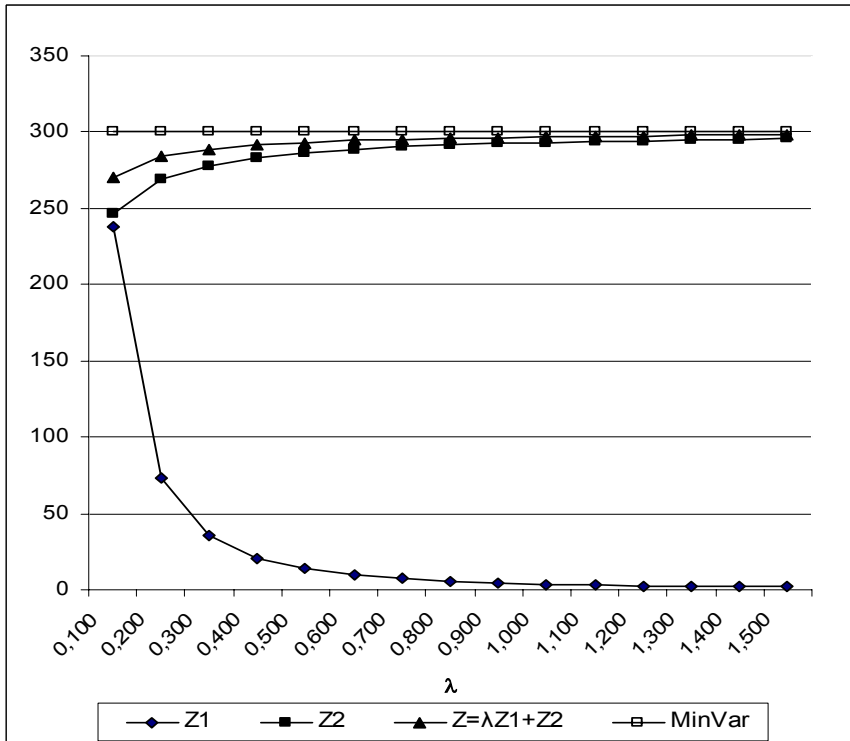
E_r ist dabei die Einheitsmatrix der Ordnung r . Da die zu invertierende Matrix hier stets regulär ist (Siehe Abschnitt 7.6 im Anhang), existiert diese Lösung grundsätzlich, was das Verfahren, das dem Ridge-Regression-Ansatz entspricht, besonders robust macht.

Dies ist ein Vorteil gegenüber dem MinVar - Verfahren. Hinweise zur Ridge-Regression findet man bei Hartung und Elpelt (1986). Allerdings können auch hier negative Gewichte nicht ausgeschlossen werden. Man erkennt übrigens in (5), dass die Lösung mit wachsendem Kontrollparameter λ gegen diejenige des MinVar - Verfahrens (3) strebt.

Chambers (1996) und Rao & Singh (1997) benutzen eine zu (4) ähnliche Zielfunktion, wobei der Abstand zwischen Gewichten und Vorgewichten mittels χ^2 -Distanzmaß gemessen wird.

Abbildung 1 zeigt einen typischen Verlauf (aus einem Beispiel mit $n = 2502$, $r = 44$) der Größen Z_1 und Z_2 im jeweiligen Minimum von $Z = \lambda Z_1 + Z_2$. Zusätzlich ist der Wert der Zielfunktion von MinVar eingetragen (ca. 300). Man erkennt, dass für λ etwa oberhalb 0,9 praktisch kein Unterschied mehr zwischen beiden Methoden vorliegt: Der Anpassungsfehler Z_1 ist praktisch 0 (was der Bedingung $Sg = b$ bei MinVar entspricht). Bei $\lambda = 0,1$ liegt der quadrierte Abstand zu den Vorgewichten Z_2 beim Alternativmodell bei ca. 250, also um 50 Einheiten besser als der Zielfunktionswert von MinVar (300).

Bei weiter fallenden Werten für λ würden sich die Gewichte bei Min I den Vorgewichten nähern (Z_2 geht gegen 0) – bei gleichzeitiger Verschlechterung der Merkmalsanpassung Z_1 . In der Praxis hat sich gezeigt, dass mit $\lambda = 1$ sehr gute Ergebnisse erzielt werden, die mit MinVar durchaus konkurrieren können. Eine weitere Steigerung über 1 hinaus würde in diesem Beispiel keine wesentliche Verbesserung der Anpassung $Z_1 = (Sg - b)^T(Sg - b)$ mehr bringen.

Abbildung 1 Einfluss von λ auf die Gewichtung

Die Matrix $E_r / \lambda + SS^T$ in (5) ist grundsätzlich regulär (Siehe Anhang 7.6). Allerdings könnten numerische Probleme bei der Matrizeninversion auftreten im Falle extrem hoher Werte r an Restriktionen.

Dieser Fall liegt dann vor, wenn nicht nur Randverteilungen von Anpassungsmerkmalen berücksichtigt werden sollen, sondern auch beliebige Kombinationen zwischen solchen Variablen zu einer erheblichen Anzahl r von Gewichtungszellen führen.

Wir leiten nun aus diesem Ansatz (4) eine Methode ab, die a) unabhängig von der Anzahl der Restriktionen numerisch stabil bleibt und b) grundsätzlich zu positiven Gewichten führt.

5 Ein robustes Gewichtungungsverfahren mit positiven Gewichten („Min II“)

Es sei nun f eine Funktion, die jedem reellen x eine reelle Zahl $f(x)$ zuordnet. Die existierenden Ableitungen $f'(x)$ und $f''(x)$ seien außerdem stetig. Wir vereinbaren folgende Schreibweisen: $f(x) = (f(x_1), \dots, f(x_n))^T$, falls das Argument x eine Vektorgröße ist. Dies gelte auch für die Funktionen $f'(x)$ und $f''(x)$.

Wir ersetzen nun die Gewichte g_i durch den Ansatz $f(\gamma_i) = g_i$ mit noch unbekannten reellen Größen γ_i und setzen $f(\gamma) = (f(\gamma_1), \dots, f(\gamma_n))^T$.

Bei einem Spaltenvektor x bedeute die Schreibweise $\text{Diag } \{x\}$ die Diagonalmatrix der Komponenten von x .

Die Zielfunktion (4) lautet dann

$$Z = \lambda (S f(\gamma) - b)^T (S f(\gamma) - b) + (f(\gamma) - t)^T (f(\gamma) - t) \quad (6)$$

Die Transformation f dient ausschließlich dazu, dass die Gewichte $f(\gamma_i)$ in noch vorzugebenden Bereichen liegen. So kann die Bedingung $f(\gamma_i) \in [a, b]$ etwa durch $f(x) = a + (b-a)\cos^2 x$ erfüllt werden. Die Funktion $f(x) = e^x$ führt zu positiven Gewichten und nicht negative Gewichte ergibt $f(x) = x^2$. Keine Einschränkung der Gewichte liefert $f(x) = x$ und (6) ist dann mit der Zielfunktion (4) identisch.

5.1 Lösung des Optimierungsproblems

Die Zielfunktion (6) lässt sich umformen in

$$Z = f(\gamma)^T [\lambda S^T S + E_n] f(\gamma) - 2[\lambda b^T S + t^T] f(\gamma) + \lambda b^T b + t^T t \quad (7)$$

Dabei ist E_n Einheitsmatrix der Ordnung n , Z im Allgemeinen keine quadratische Form in γ .

Der Gradient $\delta = \text{Grad } Z(\gamma_1, \dots, \gamma_n)$ bezüglich $\gamma_1, \dots, \gamma_n$ lautet:

$$\delta = 2 \text{Diag } \{f'(\gamma)\} ([\lambda S^T S + E_n] f(\gamma) - \lambda S^T b - t) \quad (8)$$

Zur Lösung der Optimierungsaufgabe $Z(\gamma_1, \dots, \gamma_n) = \text{Min}!$ verwenden wir das Gradientenverfahren 1. Ordnung $\gamma = \gamma - \varepsilon \delta$ mit der positiven Schrittweite ε .

5.2 Positive Gewichte: $f(\mathbf{x}) = \mathbf{e}^{\mathbf{x}}$

Wir behandeln nun den Fall $f(\mathbf{x}) = \mathbf{e}^{\mathbf{x}}$, der die Positivität der Gewichte garantiert.

Mit der Bezeichnung $\mathbf{e}^{\gamma} = (e^{\gamma_1}, \dots, e^{\gamma_n})^T$ lautet die Zielfunktion

$$Z = \mathbf{e}^{\gamma^T} [\lambda \mathbf{S}^T \mathbf{S} + \mathbf{E}_n] \mathbf{e}^{\gamma} - 2[\lambda \mathbf{b}^T \mathbf{S} + \mathbf{t}^T] \mathbf{e}^{\gamma} + \lambda \mathbf{b}^T \mathbf{b} + \mathbf{t}^T \mathbf{t} \quad (9)$$

mit dem Gradienten

$$\delta = 2 \text{Diag} \{ \mathbf{e}^{\gamma} \} ([\lambda \mathbf{S}^T \mathbf{S} + \mathbf{E}_n] \mathbf{e}^{\gamma} - \lambda \mathbf{S}^T \mathbf{b} - \mathbf{t}). \quad (10)$$

Für die Jacobische Funktionalmatrix ergibt sich (siehe Anhang)

$$J = 2 \text{Diag} \{ \mathbf{e}^{\gamma} \} [\lambda \mathbf{S}^T \mathbf{S} + \mathbf{E}_n] \text{Diag} \{ \mathbf{e}^{\gamma} \} + \text{Diag} \{ \delta \}. \quad (11)$$

5.3 Die Lösung mittels Gradientenverfahren 1. Ordnung

Mit obigen Ergebnissen (10) und (11) ergibt sich nun das Iterationsverfahren $\gamma := \gamma - \varepsilon \delta$, woraus sich die positiven Gewichte unmittelbar als $g_i = e^{\gamma_i}$ ergeben. Falls die Gewichte aus (5) ohnehin positiv waren, so stimmen sie mit den hier erhaltenen überein (Siehe den Hinweis im Anhang (7.3)). In diesem Fall macht ein Start mit Min II keinen Sinn – es sei denn, die Matrizeninversion in (5) wird wegen extrem hoher Anzahl r an Restriktionen problematisch.

Nur in diesem Ausnahmefall wird als Startwert γ beim Gradientenverfahren der Logarithmus der Vorgewichte, $\gamma_i = \ln(t_i)$, verwendet.

Falls die Min I - Lösung $\mathbf{g} = \mathbf{t} + \mathbf{S}^T (\mathbf{E}_r / \lambda + \mathbf{S} \mathbf{S}^T)^{-1} [\mathbf{b} - \mathbf{S} \mathbf{t}]$ nicht nur positive Komponenten hat, werden diejenigen Werte, die unterhalb einer Schranke (etwa 0,001) liegen, auf diesen positiven Grenzwert gesetzt und das Gradientenverfahren mit dem Logarithmus der so veränderten Gewichte gestartet. Mit der Schrittweite $\varepsilon = \delta^T \delta / \delta^T J \delta$ konvergiert das Verfahren dann sehr rasch.

Da in der Praxis Gewichtungen mit über $n = 50000$ Befragten möglich sind, ist der Hinweis wichtig, dass bei der Softwareimplementierung von Min II das Gradientenverfahren nur drei Arbeitsfelder der Größe n benötigt. Gradientenverfahren quadratischer Konvergenz (Newton-Raphson, Konjugierte Gradienten) scheitern daran, dass die Inverse einer (n,n) -Matrix berechnet werden müsste.

Das einfache Gradientenverfahren hat – als Iterationsverfahren – den Vorteil gegenüber nicht iterativen Methoden, dass es einerseits selbstkorrigierend ist (Rundungsfehler kön-

nen sich nicht „aufschaukeln“) und außerordentlich numerisch stabil ist im Vergleich zu anderen denkbaren Lösungsansätzen. Dieser Effekt ist gerade bei sehr hohen Fallzahlen n von entscheidender Bedeutung.

Fehlende Werte (Missing Values) bei einer bestimmten Kategorie machen sich dadurch bemerkbar, dass eine der r Zeilen der Matrix S zur Nullzeile wird. Dies führt – wie anfangs erwähnt – zu Problemen bei MinVar.

Beim Alternativmodell Min I spielt dies offensichtlich rechnerisch keine Rolle, weil dort in (5) die Matrix $E_r / \lambda + SS^T$ grundsätzlich regulär bleibt (Siehe Anhang 7.6). Gleiches gilt natürlich für Min II.

6 Rechenbeispiele aus der Praxis

6.1 Das Datenmaterial

Für ein Forschungsprojekt am Institut für Markt-, Media- und Regionalforschung BIK ASCHPURWIS+BEHRENS, Hamburg, wurde eine kleinere face to face-Befragung von 5007 Befragten (ab 14 Jahre) an Sollvorgaben einer repräsentativen face to face-Bevölkerungsumfrage mit insgesamt 32423 Interviews angepasst. Zur Demonstration werden wir diese nun bezüglich der Merkmale Alter (7 Kat.), Geschlecht (2), Schule (höchster Schulabschluss des Befragten, die letzte Ausprägung „Schüler“ sind die Befragten, die noch zur Schule gehen.) (4), Haushaltsgröße (4) und Nutzergruppe eines Mediums nach Leseorten (9) an die größere Befragung anpassen. Beide Studien weisen bezüglich dieser Variablen vergleichbare Fragestellungen auf. Als Vorgewichte verwenden wir in diesem Beispiel einheitlich Einsen.

Die Schwierigkeiten liegen in den Diskrepanzen bei dem Merkmal Nutzergruppe, wie Tabelle 1 zeigt, weswegen wir dieses Beispiel zu Demonstrationszwecken bewusst gewählt haben.

Tabelle 1 Nutz-GRP

| Kategorie | Sollwerte | Istwerte vor Gewichtung |
|-----------|-----------|-------------------------|
| 0 | 1511 | 2603 |
| 1 | 733 | 448 |
| 2 | 2480 | 1881 |
| 10 | 17 | 8 |
| 11 | 44 | 4 |
| 12 | 78 | 18 |
| 20 | 14 | 7 |
| 21 | 35 | 13 |
| 22 | 95 | 25 |

6.2 Beispiel 1

Mit dem anfänglichen Kontrollparameter $\lambda = 1$ wurden zunächst die Methoden Min I und Min II gerechnet; zum Vergleich wurden auch die Resultate von MinVar und IPF mitgeliefert.

Tabelle 2 zeigt die Unterschiede der 4 beschriebenen Gewichtungungsverfahren. IPF hat tendenziell höhere Gewichte bei guter Genauigkeit der Anpassungsmerkmale.

Allerdings geht dies zu Lasten des größeren Abstands zu den Vorgewichten.

MinVar und Min I liefern in diesem Beispiel negative Gewichte, wobei erwartungsgemäß MinVar die Struktur der Merkmale fehlerfrei reproduziert wegen der Bedingung (1) $Sg=b$. Am nächsten an den Vorgewichten liegen die Ergebnisse von Min I und Min II mit gleicher Güte bezüglich der Anpassungsgenauigkeit.

Die Effektivität ist ein Maß für die Streuung der Gewichte, definiert als Quotient aus quadrierter Gewichtssumme zur Summe der Quadrate der Gewichte (prozentuiert auf die Anzahl n der Befragten). Hohe Werte sind gleichbedeutend mit niedriger Streuung. So ist die Höhe der Effektivität ein Qualitätsmerkmal für die erhaltene Gewichtung.

Nur dann, wenn alle Gewichte einander gleich wären (Streuung 0), wäre die Effektivität gleich 100 %. Aus Tab. 2 ist ersichtlich, dass IPF die geringste Effektivität hat. Am höchsten liegt sie bei Min I und Min II.

Tabelle 2 Kenngrößen Gewichte

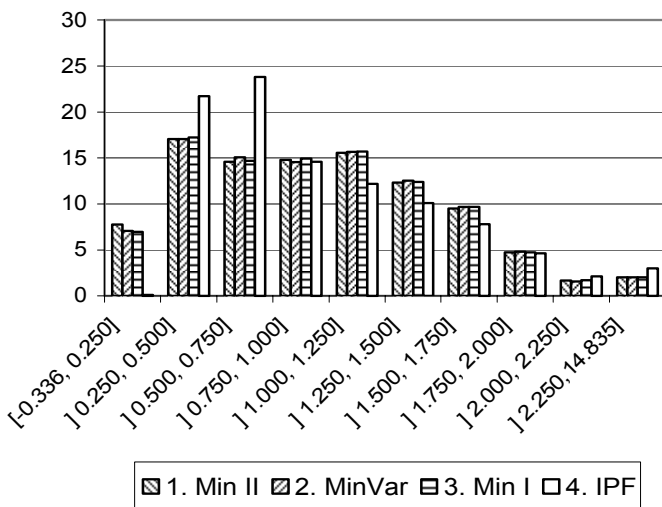
| | Effektivität | Minimum g | Max g | Quadrierter Abstand zu den Vorgewichten | Anpassungsgenauigkeit (= Quadrierter Abstand S g zu b) |
|---------------|--------------|--------------|----------|---|---|
| Min II | 70,30% | 0,017 | 9,673 | 2117,4 | 58,3 |
| MinVar | 69,05% | -0,336 | 11,235 | 2243,7 | 0 |
| Min I | 70,30% | -0,325 | 9,67 | 2117,1 | 58,2 |
| IPF | 67,13% | 0,207 | 14,835 | 2451,3 | 0,1 |

Dass diese Unterschiede der Anpassungsgenauigkeit bei allen 4 Methoden indessen nicht sehr voneinander abweichen, zeigt Tabelle 3 mit den Soll-Ist-Abweichungen der errechneten Fallzahlen. Erwartungsgemäß ist der Unterschied bei Kategorie 11 von Nutz-GRP am stärksten (Sollwert 44, heraus kommt ein Istwert von 38 gewichteten Fällen bei Min I, Min II für $\lambda = 1$).

Tabelle 3 Abweichungen

| $(\lambda=1)$ | | | | | | |
|-------------------|-----------|------|--------------------------------|-------------|-------------------------------|----------|
| Merkmal | Kategorie | Soll | Soll-Min II ($\lambda=1$) | Soll-MinVar | Soll-Min I ($\lambda=1$) | Soll-IPF |
| Alter | 14-19 J | 389 | 0 | 0 | 0 | 0 |
| | 20-29 J | 591 | 0 | 0 | 0 | 0 |
| | 30-39 J | 884 | 0 | 0 | 0 | 0 |
| | 40-49 J | 882 | 0 | 0 | 0 | 0 |
| | 50-59 J | 723 | 0 | 0 | 0 | 0 |
| | 60-69 J | 805 | 0 | 0 | 0 | 0 |
| | 70 J + | 734 | 1 | 0 | 1 | 0 |
| Geschlecht | Männer | 2398 | 1 | 0 | 1 | 0 |
| | Frauen | 2609 | 1 | 0 | 1 | 0 |
| Schule | Haupt | 2288 | 0 | 0 | 0 | 0 |
| | Mittel | 1630 | 0 | 0 | 0 | 0 |
| | Abitur | 805 | 0 | 0 | 0 | 0 |
| | Schüler | 284 | 2 | 0 | 1 | 0 |
| HHGR | 1 Person | 996 | 0 | 0 | 0 | 0 |
| | 2 Pers. | 1847 | 0 | 0 | 0 | 0 |
| | 3 Pers. | 967 | 1 | 0 | 1 | 0 |
| | 4 Pers. + | 1197 | 1 | 0 | 1 | 0 |
| Nutz-GRP | 0 | 1511 | -2 | 0 | -2 | 0 |
| | 1 | 733 | -1 | 0 | -1 | 0 |
| | 2 | 2480 | -2 | 0 | -2 | 0 |
| | 10 | 17 | -1 | 0 | -1 | 0 |
| | 11 | 44 | 6 | 0 | 6 | 0 |
| | 12 | 78 | 1 | 0 | 1 | 0 |
| | 20 | 14 | -1 | 0 | -1 | 0 |
| | 21 | 35 | 0 | 0 | 0 | 0 |
| | 22 | 95 | 1 | 0 | 1 | 0 |

Das Gros der Gewichte liegt unterhalb 2,250, wie aus der Abb. 2 hervorgeht. Bis auf IPF liefern alle 4 Methoden in etwa die gleiche Gewichtsverteilung. Die IPF-Gewichte häufen sich zwischen 0,250 und 0,750.

Abbildung 2 Verteilung der Gewichte

Mit wachsendem Kontrollparameter λ erhöht sich die Anpassungsgüte der Merkmale an die Sollvorgaben. In Tabelle 4 geht der Fehler bei Min I und Min II auf rund 17 zurück, wenn λ von ursprünglich 1 auf 2 ansteigt. Im „Problemfall“ Kategorie 11 der Variablen Nutz-GRP gilt nun Soll = 44, Ist = 41 Fälle.

Tabelle 4 Abweichung

| ($\lambda=2$) | | | | | | |
|---------------------------------|------------------|-------------|---|-------------------------|--|-----------------|
| Merkmal | Kategorie | Soll | Soll-Min II ($\lambda=2$) | Soll- MinVar | Soll-Min I ($\lambda=2$) | Soll-IPF |
| Alter | 14-19 J | 389 | 0 | 0 | 0 | 0 |
| | 20-29 J | 591 | 0 | 0 | 0 | 0 |
| | 30-39 J | 884 | 0 | 0 | 0 | 0 |
| | 40-49 J | 882 | 0 | 0 | 0 | 0 |
| | 50-59 J | 723 | 0 | 0 | 0 | 0 |
| | 60-69 J | 805 | 0 | 0 | 0 | 0 |
| | 70 J+ | 734 | 0 | 0 | 0 | 0 |
| Geschlecht | Männer | 2398 | 1 | 0 | 1 | 0 |
| | Frauen | 2609 | 0 | 0 | 0 | 0 |
| Schule | Haupt | 2288 | 0 | 0 | 0 | 0 |
| | Mittel | 1630 | 0 | 0 | 0 | 0 |
| | Abitur | 805 | 0 | 0 | 0 | 0 |
| | Schüler | 284 | 1 | 0 | 1 | 0 |
| HHGR | 1 Person | 996 | 0 | 0 | 0 | 0 |
| | 2 Pers. | 1847 | 0 | 0 | 0 | 0 |
| | 3 Pers. | 967 | 0 | 0 | 0 | 0 |
| | 4 Pers. + | 1197 | 0 | 0 | 0 | 0 |
| Nutz-GRP | 0 | 1511 | -1 | 0 | -1 | 0 |
| | 1 | 733 | -1 | 0 | -1 | 0 |
| | 2 | 2480 | -1 | 0 | -1 | 0 |
| | 10 | 17 | 0 | 0 | 0 | 0 |
| | 11 | 44 | 3 | 0 | 3 | 0 |
| | 12 | 78 | 1 | 0 | 1 | 0 |
| | 20 | 14 | 0 | 0 | 0 | 0 |
| | 21 | 35 | 0 | 0 | 0 | 0 |
| | 22 | 95 | 0 | 0 | 0 | 0 |
| Effektivität | | | 69,73% | 69,05% | 69,73% | 67,13% |
| Min g | | | 0,009 | -0,336 | -0,331 | 0,207 |
| Max g | | | 10,378 | 11,235 | 10,373 | 14,835 |
| Anpassungsfehler | | | 16,989 | 0 | 16,963 | 0,1 |
| Quadr. Abst. Vorgew. | | | 2173,17 | 2243,689 | 2172,755 | 2451,3 |

Wird der Kontrollparameter λ auf 5 gesetzt, so ergibt sich beim Merkmal Nutz-GRP folgendes Bild (bei allen anderen Merkmalen gibt es keine Unterschiede mehr):

Tabelle 5

| $(\lambda=5)$ | | | | | | |
|------------------------------|-----------|------|-------------|-------------|------------|----------|
| Merkmal | Kategorie | Soll | Soll-Min II | Soll-MinVar | Soll-Min I | Soll-IPF |
| Nutz-GRP | 0 | 1511 | 0 | 0 | 0 | 0 |
| | 1 | 733 | 0 | 0 | 0 | 0 |
| | 2 | 2480 | 0 | 0 | 0 | 0 |
| | 10 | 17 | 0 | 0 | 0 | 0 |
| | 11 | 44 | 1 | 0 | 1 | 0 |
| | 12 | 78 | 0 | 0 | 0 | 0 |
| | 20 | 14 | 0 | 0 | 0 | 0 |
| | 21 | 35 | 0 | 0 | 0 | 0 |
| | 22 | 95 | 0 | 0 | 0 | 0 |
| | | | | | | |
| Effektivität | | | 69,34% | 69,05% | 69,34% | 67,13% |
| Min g | | | 0,007 | -0,336 | -0,331 | 0,207 |
| Max g | | | 10,874 | 11,235 | 10,868 | 14,835 |
| Anpassungsfehler | | | 3,01 | 0 | 3 | 0,1 |
| Quadr. Abst. zu Vorgewichten | | | 2213,5 | 2243,689 | 2213,1 | 2451,3 |

Nach wie vor liegen die Gewichte bei Min I und Min II näher an den Vorgewichten. Der einzige Unterschied liegt bei Kategorie 11: Statt 44 Fällen ergeben sich nun 43 gewichtete Fälle bei beiden Verfahren (ab $\lambda = 6$ ist kein Unterschied bei der Anpassung mehr vorhanden und der quadratische Abstand zu den Vorgewichten liegt mit 2223,7 für Min II bzw. 2223,5 für Min I immer noch besser als bei MinVar und erst recht IPF mit 2451,3).

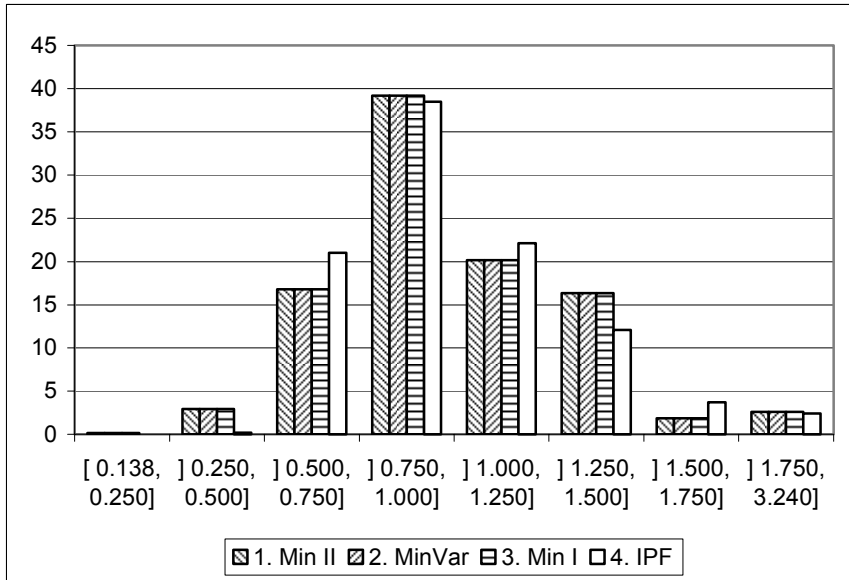
6.3 Beispiel 2

Verzichtet man auf das problematische Merkmal Nutz-GRP (vgl. Tab. 1), bleiben die Gewichte bei allen 4 Methoden positiv. In diesem Fall liefern Min I und Min II (bei $\lambda = 1$) dieselben Ergebnisse, wie abschließend Tab. 6 zeigt. Min II bricht nach der ersten Iteration ab.

Auch hier wird wieder die Sonderstellung von IPF deutlich: Die Gewichte weichen deutlich von denen der übrigen Verfahren ab (Abb. 3). Dagegen stimmen die MinVar-Gewichte mit denjenigen von Min I, Min II praktisch überein.

Tabelle 6

| $(\lambda = 1)$ | | | | | | |
|---------------------------------|-----------|------|----------------|-----------------|------------|----------|
| Merkmal | Kategorie | Soll | Soll-Min II | Soll- MinVar | Soll-Min I | Soll-IPF |
| Alter | 14-19 J | 389 | 0 | 0 | 0 | 0 |
| | 20-29 J | 591 | 0 | 0 | 0 | 0 |
| | 30-39 J | 884 | 0 | 0 | 0 | 0 |
| | 40-49 J | 882 | 0 | 0 | 0 | 0 |
| | 50-59 J | 723 | 0 | 0 | 0 | 0 |
| | 60-69 J | 805 | 0 | 0 | 0 | 0 |
| | 70 J+ | 734 | 0 | 0 | 0 | 0 |
| Geschlecht | Männer | 2398 | 0 | 0 | 0 | 0 |
| | Frauen | 2609 | 0 | 0 | 0 | 0 |
| Schule | Haupt | 2288 | 0 | 0 | 0 | 0 |
| | Mittel | 1630 | 0 | 0 | 0 | 0 |
| | Abitur | 805 | 0 | 0 | 0 | 0 |
| | Schüler | 284 | 1 | 0 | 1 | 0 |
| HHGR | 1 Person | 996 | 0 | 0 | 0 | 0 |
| | 2 Pers. | 1847 | 0 | 0 | 0 | 0 |
| | 3 Pers. | 967 | 0 | 0 | 0 | 0 |
| | 4 Pers.+ | 1197 | 0 | 0 | 0 | 0 |
| | | | Min II | MinVar | Min I | IPF |
| Effektivität | | | 90,05 % | 89,99% | 90,05% | 89,86% |
| min g | | | 0,148 | 0,138 | 0,148 | 0,388 |
| max g | | | 2,589 | 2,606 | 2,589 | 3,24 |
| Quadr. Abst. Vorgewichte | | | 553,286 | 556,956 | 553,286 | 565,286 |
| Anpassungsgenauigkeit | | | 1,825 | 0 | 1,825 | 0,643 |

Abbildung 3 Verteilung der Gewichte

6.4 Ein Experiment mit zusammengesetzten Merkmalen

Zum Abschluss folgt noch ein Extrembeispiel mit untereinander verknüpften Merkmalen.

In Tabelle 7 sind 8 Merkmale aufgeführt, von denen die letzten 4 aus den ersten 4 zusammengesetzt sind. Es handelt sich hier um insgesamt $r = 327$ Restriktionen (bei $n = 5007$ Befragten).

In diesem Fall versagt IPF wegen der entstehenden 15 leeren Ist-Zellen, was allenfalls durch ein recht umständliches Zusammenlegen von Kategorien verhindert werden könnte.

Auch MinVar funktioniert hier nicht mehr, weil a) 55 Zeilen in S zu Null werden und b) die kombinierten Variablen 5 bis 8 zu den erwähnten Redundanzen im System der Restriktionen führt.

Erwartungsgemäß ergeben sich keine Probleme bei Min II. Mit $\lambda = 1$ ergeben sich sehr gute Ergebnisse, wie die Tabelle 7 zeigt.

Die höchste Abweichung zwischen Soll und Ist ergibt sich bei den Frauen über 70 Jahren mit Hauptschulabschluss in einem Haushalt mit mehr als 4 Personen: Statt der 11 Sollfälle liefert die Min II - Gewichtung 8 solcher Fälle. In allen anderen der 327 Ausprägungen liegt der Unterschied unter 0,1 % bezogen auf 5007 Fälle.

Tabelle 7

| | | Leere Soll-Zellen | leere Ist-Zellen | Nur bei Ist leere Zellen |
|--|--------|----------------------|---------------------|-----------------------------|
| Merkmale | | | | |
| Alter (7) | | | | |
| Sex (2) | | | | |
| Schule (4) | | | | |
| HHGR (4) | | | | |
| Alter *Sex (14) | | | | |
| Schule *HHGR (16) | | | | |
| Alter * Sex * HHGR (56) | | | | |
| Alter * Sex * Schule *HHGR (224) | | 40 | 55 | 15 |
| Min g | 0,143 | | | |
| Max g | 7,573 | | | |
| Effektivität | 82,15% | | | |
| Quadr. Abst. Vorgewichte | 1087,7 | | | |
| Anpassungs- güte bei 327 Restriktionen | 68,43 | | | |
| Höchster Soll-Ist-Unterschied bei Frauen, 70 J+, Hauptschule, 4 Pers. + : Soll = 11, Ist = 8 Fälle Alle übrigen Unterschiede liegen – bezogen auf 5007 Fälle – unter 0,1 %. | | | | |

7 Anhang

7.1 Spezielle Formeln für Gradienten und Jacobi-Matrix

Unter der Jacobischen Funktionalmatrix J einer reellwertigen Funktion $F(x_1, \dots, x_n)$ verstehen wir die (n, n) -Matrix der partiellen Ableitungen $\partial^2 F(x_1, \dots, x_n) / (\partial x_i \partial x_k)$. Es wird dabei unterstellt, dass die partiellen Ableitungen existieren und von der Reihenfolge unabhängig sind.

Wir haben in den vorangegangenen Abschnitten einige Formeln gebraucht, die hier kurz erläutert werden sollen. Sie betreffen Gradienten und Jacobische Funktionalmatrix bezüg-

lich γ_i der reellen Funktion $f(\gamma)^T A f(\gamma)$ mit symmetrischer Matrix A und der Funktion $b^T f(\gamma)$. Dabei hängen A und der Vektor b nicht von den γ_i ab. Nicht indizierte Größen sind Spaltenvektoren. Wie schon erwähnt, sind $f'(\gamma_i)$ und $f''(\gamma_i)$ die Werte der 1. und 2. Ableitung der Funktion f an der Stelle γ_i , $f'(\gamma)$ und $f''(\gamma)$ die entsprechenden Spaltenvektoren.

Ist v ein Vektor (in Spaltendarstellung), so bezeichne $\text{Diag} \{v\}$ die Diagonalmatrix aus den Komponenten dieses Vektors. Bei besonderem Bezug auf die Komponenten v_i des Vektors v gelte auch Schreibweise $\text{Diag} \{v_i\}$. Es folgt:

$$\begin{aligned}\text{Grad } f(\gamma)^T A f(\gamma) &= 2 \text{Diag} \{f'(\gamma)\}^T A f(\gamma) \\ \text{Grad } b^T f(\gamma) &= \text{Diag} \{f'(\gamma)\}^T b\end{aligned}$$

Funktionalmatrix von $b^T f(\gamma) : \text{Diag} \{b_i f''(\gamma_i)\}$,
und von $f(\gamma)^T A f(\gamma) : 2 \text{Diag} \{f'(\gamma)\}^T A \text{Diag} \{f'(\gamma)\} + 2 \text{Diag} \{f''(\gamma)\}^T A f(\gamma)$

7.2 Herleitung Formel (3)

Wir minimieren die Lagrange-Funktion $L(g, \mu) = (g-t)^T(g-t) + \mu^T(Sg-b)$, wobei die Minimierung bezüglich g und des r -dimensionalen Parametervektors μ erfolgt. Für die notwendigen Bedingungen setzen wir die Gradienten auf Null.

$$\begin{aligned}\text{Grad}_g L &= 2(g-t) + S^T \mu = 0 & (*) \\ \text{Grad}_\mu L &= Sg-b = 0 & (**)\end{aligned}$$

Aus (*) ergibt sich $S^T \mu = -2(g-t)$. Multiplikation mit S von links liefert zunächst $SS^T \mu = -2S(g-t)$. Falls S maximalen Rang hat, folgt $\mu = -2(SS^T)^{-1}S(g-t)$. Also unter Verwendung von (**): $\mu = -2(SS^T)^{-1}S(g-t) = -2(SS^T)^{-1}(b-St)$.

Daraus folgt zusammen mit (*) schließlich (3). Falls S nicht maximalen Rang hat, kann man versuchen, mit Hilfe der Moore-Penrose-Inversen von SS^T zu einer Ersatzlösung zu kommen.

In Hartung und Elpelt (1986) findet man einen entsprechenden Algorithmus zur Bildung der Pseudoinversen.

7.3 Herleitung Formel (5)

Wir setzen den Gradienten von $Z = \lambda(Sg - b)^T(Sg - b) + (g-t)^T(g-t)$ gleich 0 und erhalten $(\lambda S^T S + E_n)g = \lambda S^T b + t$. Es folgt nach Linksmultiplikation mit S $(\lambda SS^T S + S)g = \lambda S S^T b + St$ und damit $(\lambda SS^T + E_r) S g = \lambda S S^T b + St$, also $S g = (\lambda SS^T + E_r)^{-1}(\lambda S S^T b + St)$.

Aus der ursprünglichen Beziehung $(\lambda S^T S + E_n)g = \lambda S^T b + t$ ergibt sich $g = -\lambda S^T S g + \lambda S^T b + t$. Darin ersetzen wir nun $S g$ durch $(\lambda SS^T + E_r)^{-1}(\lambda S S^T b + St)$:

$$\begin{aligned} g &= t + \lambda S^T b - \lambda S^T (\lambda SS^T + E_r)^{-1} (\lambda S S^T b + St) = t + \lambda S^T b - S^T (SS^T + E_r/\lambda)^{-1} (\lambda S S^T b + St), \\ g &= t + \lambda S^T b - S^T (SS^T + E_r/\lambda)^{-1} (-b + b + \lambda S S^T b + St), \\ g &= t + \lambda S^T b - S^T (SS^T + E_r/\lambda)^{-1} (E_r + \lambda S S^T) b - S^T (SS^T + E_r/\lambda)^{-1} (-b + St), \\ g &= t + \lambda S^T b - \lambda S^T b - S^T (SS^T + E_r/\lambda)^{-1} (-b + St) = t + S^T (E_r/\lambda + SS^T)^{-1} [b - St]. \end{aligned}$$

Hinweis: Sind alle Komponenten von g positiv, so verschwindet für dieses g mit $\gamma := \text{Ln}(g)$ der Gradient in (10), $\delta = 2 \text{Diag}\{e^\gamma\} ([\lambda S^T S + E_n] e^\gamma - \lambda S^T b - t)$; denn es gilt ja dann:

$$[\lambda S^T S + E_n] e^\gamma - \lambda S^T b - t = [\lambda S^T S + E_n] g - \lambda S^T b - t.$$

Und dies ist gleich 0 nach obiger Herleitung. Diese Überlegung gilt natürlich entsprechend für alle anderen Funktionen $f(x)$, wenn man allgemein $\gamma := f^{-1}(g)$ setzt mit f^{-1} als Umkehrfunktion von f .

7.4 Formel (11)

Für die Jacobische Funktionalmatrix J von (9) ergibt sich zunächst

$$2\text{Diag}\{e^\gamma\} [\lambda S^T S + E_n] \text{Diag}\{e^\gamma\} + 2\text{Diag}\{e^\gamma\} \text{Diag}\{[\lambda S^T S + E_n] e^\gamma - \lambda S^T b - t\}.$$

Der 2. Summand aber ist aber gerade gleich $\text{Diag}\{\delta\}$:

Sei $D = 2\text{Diag}\{e^\gamma\}$ und $v = [\lambda S^T S + E_n] e^\gamma - \lambda S^T b - t$.

Nun ist wegen (10): $\delta = 2 \text{Diag}\{e^\gamma\} ([\lambda S^T S + E_n] e^\gamma - \lambda S^T b - t) = D v$.

Also $\text{Diag}\{\delta\} = \text{Diag}\{Dv\} = D \text{Diag}\{v\} = 2\text{Diag}\{e^\gamma\} \text{Diag}\{[\lambda S^T S + E_n] e^\gamma - \lambda S^T b - t\}.$

7.5 Notwendige Änderungen bei Verwendung des χ^2 – Distanzmaßes

Es bezeichne $D = \text{diag} \{t\}$ die Diagonalmatrix der Eingangsgewichte. Damit lässt sich der χ^2 – Abstand zwischen den Gewichten g und t darstellen in der Form: $(g-t)^T D^{-1} (g-t)$.

Man kann zeigen, dass sich mit diesem Distanzbegriff die bisherigen Formeln, die für den quadrierten Euklidischen Abstand gelten, so schreiben lassen:

$$\text{MinVar:} \quad g = t + DS^T(SDS^T)^{-1}[b - S t] \quad (3)$$

$$\text{Min I:} \quad g = t + DS^T(E_r / \lambda + SDS^T)^{-1} [b - S t] \quad (5)$$

$$\text{Min II:} \quad \delta = 2 \text{ Diag} \{e^\gamma\} ([\lambda S^T S + D^{-1}] e^\gamma - \lambda S^T b - D^{-1} t) \quad (10)$$

Hinweis: In der technischen Implementierung (in APL) von Min I und Min II ist obige χ^2 -Version als Option enthalten. Falls – wie in den Beispielen – alle Vorgewichte gleich 1 sind, stimmt das χ^2 - Distanzmaß natürlich mit dem quadrierten Euklidischen Abstand überein.

7.6 Hinweise zur Matrix $E_r / \lambda + SS^T$

Für die Indikatormatrix S gilt grundsätzlich: Werden alle Ausprägungen eines Anpassungsmerkmals berücksichtigt, so liefert die Summe der entsprechenden Zeilen von S nur Einsen. Das hat zur Folge, dass S automatisch eine Rangreduktion erfährt, wenn bei einem weiteren Anpassungsmerkmal alle Ausprägungen eingetragen werden. Dies führt zu den erwähnten Problemen bei dem MinVar-Verfahren.

Der Rang der (r, n) -Matrix S (wobei stets $n > r$) ist gleich dem Rang des Gaußschen Produkts SS^T . Diese symmetrische Matrix ist stets positiv semidefinit: Alle ihre Eigenwerte sind nicht negativ. Durch Addition der mit dem positiven Faktor $1/\lambda$ multiplizierten Einheitsmatrix ergibt sich für die Matrix $E_r / \lambda + SS^T$: Alle Eigenwerte sind positiv, weil zu den (nicht negativen) Eigenwerten von SS^T der positive Summand $1/\lambda$ hinzukommt. Insbesondere ist die Matrix damit stets invertierbar (Determinante = Produkt der Eigenwerte). Dieser Hinweis gilt ebenso für die Matrix $E_r / \lambda + SDS^T$ im Fall des χ^2 – Distanzmaßes (7.5).

8 Fazit – Abschließende Bemerkungen

Die Überlegungen haben gezeigt, dass Min I und seine Erweiterung auf positive Gewichte Min II sich durch ihre Robustheit auszeichnen bezüglich der anzupassenden Daten.

Dies ist ein entscheidender Vorteil gegenüber MinVar, das bei Redundanzen aber auch fehlenden Werten in den Daten nicht mehr richtig arbeiten kann.

Wichtig ist der Hinweis, dass Min II weder durch die Anzahl der Befragten n noch durch die der Restriktionen praktisch eingeschränkt wird.

Die gezeigten Beispiele veranschaulichen die Möglichkeit von Min II, Bereiche für die Gewichte einzuhalten (z.B. Positivitätsbedingung), was so bei MinVar nicht ohne weiteres möglich ist. Durch Variation der Kontrollgröße λ lässt sich bei Min I und Min II eine hohe Flexibilität erreichen.

Ein wachsender Parameter λ führt dazu, dass das System der Restriktionen beliebig genau erfüllt werden kann; sinkende Parameterwerte begünstigen die Nähe zu den Vorgewichten. In der Praxis hat sich jedoch der Wert $\lambda = 1$ als Richtgröße bewährt.

Die Untersuchung hat außerdem gezeigt, dass die Ergebnisse des IPF- Algorithmus – in seiner klassischen Form – von denen der übrigen vorgestellten Gewichtungungsverfahren am meisten abweichen. Die Effektivitäten der erhaltenen Gewichte sind hier am niedrigsten.

Die Problematik von IPF bezüglich leerer Ist-Zellen (Soll / Ist) spielt bei Min I und Min II keine Rolle, weil hier diese Quotientenbildung unterbleibt. Dieses Problem gilt auch für MinVar, weil dann die dortige Matrix nicht mehr invertierbar ist.

Zum Abschluss möchte ich mich besonders bedanken bei Frau Christiane Heckel, BIK-Aschpurwis + Behrens, Hamburg, die die Daten für die Beispiele zur Verfügung gestellt hat und mir wichtige Anregungen und Zusatzinformationen gegeben hat.

Literatur

- Chambers, R. L. (1996). Robust Case-Weighting for Multipurpose Establishment Surveys. *Journal of Official Statistics*, 12(1), 3-32.
- Deming, E. & Stephan, F. (1940). On a Least Square Adjustment of a Sampled Frequency Table when Expected Marginal Totals are known. *The Annals of Mathematical Statistics*, 11, 427-444.
- Hartung, J. & Elpelt B. (1986). *Multivariate Statistik. Lehr- und Handbuch der angewandten Statistik*. München, Wien: Oldenbourg.

- Merz, J. (1993). *ADJUST. Ein Programmpaket zur Hochrechnung von Mikrodaten nach dem Prinzip des minimalen Informationsverlustes*. Programm-Handbuch. FFB-Dokumentation Nr. 1.
- Merz, J. (2004). Kumulation von Mikrodaten – Konzeptionelle Grundlagen und ein Vorschlag zur Realisierung. *Allgemeines Statistisches Archiv*, 88, 451-472.
- Rao, J. N. K. & Singh, A. C. (1997). A Ridge-Shrinkage Method for Range Restricted Weight Calibration in Survey Sampling. In *Proceedings of the Section on Survey Research Methods* (pp. 57-65). Alexandria, VA: American Statistical Association.
- Rösch, G. (1994). Kriterien der Gewichtung einer nationalen Bevölkerungsstichprobe. In: S. Gabler et al. (Eds.), *Gewichtung in der Umfragepraxis* (pp. 7-26). Opladen: Westdeutscher Verlag.
- Von der Heyde, C. (1999). Techniken und Möglichkeiten der Realisierung von Random-Stichproben. In: ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V./AGMA Arbeitsgemeinschaft Media-Analyse e.V. (Eds.), *Stichproben-Verfahren in der Umfrageforschung. Eine Darstellung für die Praxis* (pp. 35-60). Opladen: Leske + Budrich.

Korrespondenzadresse

Dipl.-Math. Bruno Kopp
Bockheberer Weg 54
29640 Schneverdingen
Email: bkopp10102@aol.com